

Wright State University

**CORE Scholar**

---

Computer Science & Engineering Syllabi

College of Engineering & Computer Science

---

Winter 2012

## CS 790-02: Advanced Data Mining

Guozhu Dong

*Wright State University - Main Campus*, [guozhu.dong@wright.edu](mailto:guozhu.dong@wright.edu)

Follow this and additional works at: [https://corescholar.libraries.wright.edu/cecs\\_syllabi](https://corescholar.libraries.wright.edu/cecs_syllabi)



Part of the [Computer Engineering Commons](#), and the [Computer Sciences Commons](#)

---

### Repository Citation

Dong, G. (2012). CS 790-02: Advanced Data Mining. .  
[https://corescholar.libraries.wright.edu/cecs\\_syllabi/940](https://corescholar.libraries.wright.edu/cecs_syllabi/940)

This Syllabus is brought to you for free and open access by the College of Engineering & Computer Science at CORE Scholar. It has been accepted for inclusion in Computer Science & Engineering Syllabi by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

# CS 790-02: Advanced Data Mining

Winter, 2012

**Description:** This advanced data mining course covers concepts and techniques in data mining.

The course's focus in 2012 is on contrast data mining.

Contrast data mining has received a lot of attention since the late 1990s. Much has been done on the definition of various kinds of contrast patterns and models and various efficient mining algorithms. Much has been done on using contrasts for data mining tasks, such as applications in classification and in improving traditional classifiers, applications in rare-class/imbalanced classification, applications in outlier detection, applications in clustering. Much has been done on using contrasts and contrast-based methods to solve problems from other disciplines, including the analysis of and improvement on genetic algorithms, network security, crime prevention, privacy preservation, vision and image categorization, blog/document analysis and understanding, election and general GIS data analysis, activity recognition, microarray based disease analysis, microarray technology concordance analysis, DNA/protein site prediction, power line safety prediction, music melody analysis and trend discovery, tourism trend analysis, customer behavior change mining, rental property price prediction, prediction of diseases for patients, birth defect prediction, chemical compounds screening and selection, etc etc.

**Text book:** Contrast Data Mining: Concepts, Algorithms and Applications. Guozhu Dong and James Bailey. To appear.

The book contains 22 technical chapters on the topics discussed above. Various chapters will be distributed to the class. Students are not allowed to distribute the chapters in any manner.

**Prerequisite:** CS705 (Introduction to Data Mining), or consent of the instructor.

**Instructor:** Dr. Guozhu Dong.

**Office:** 383 Joshi

**Phone No.:** (937)-775-5066

**Email:** guozhu.dong@wright.edu

**Class time and venue:** 4:10 – 5:25, TR. RC 302.

**Office hours:** 12:50 – 1:50, TR. Use e-mail for short questions.

**Format:** This course will draw materials from the text book, and from a range of recent research papers.

Students will need to submit review reports on the papers/chapters, actively participate in class discussions, present papers, and complete a programming project.

**Problem brainstorm:** Each student needs to identify, motivate, formalize, and propose potential solutions of, a problem (related to the course materials) that the student wishes to solve. Students are not required to completely solve these problems, but they are encouraged to do so.

**Term Project:** Students will complete an implementation project, on instructor-approved topics. There will be a demo and a report, and the codes must be submitted. Students should talk to the instructor early to select paper for the project.

**Evaluation:** Final grades in the course will be determined based on students' performance on the components listed above.

**Other Resources**     • <http://www.kdnuggets.com/>

- [www.scholar.google.com](http://www.scholar.google.com)
- DBLP Bibliography Server: <http://www.informatik.uni-trier.de/~ley/db/index.html>
- ACM Digital Library: <http://www.acm.org/dl/> (Free access from WRIGHT domain.)
- IEEE Xplore: <http://ieeexplore.ieee.org/Xplore/>
- ResearchIndex (citeseer):[http: //citeseer.ist.psu.edu/](http://citeseer.ist.psu.edu/)